

### REMARKS

Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60, and 62-64 are currently pending in the subject application and are presently under consideration. Claims 1, 19, 30, 42, 44, 53, 54 and 62-64 have been amended as shown on pp. 2-20 of the Reply. Favorable reconsideration of the subject patent application is respectfully requested in view of the comments and amendments herein.

#### **I. Rejection of Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60, and 62-64 Under 35 U.S.C. §101**

Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60, and 62-64 stand rejected under 35 U.S.C. §101 as allegedly being directed to non-statutory subject matter. Withdrawal of this rejection is requested for at least the following reasons. Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60, and 62-64 produce a useful, concrete and tangible result.

Because the claimed process applies the Boolean principle [abstract idea] ***to produce a useful, concrete, tangible result*** ... on its face the claimed process comfortably falls within the scope of §101. *AT&T Corp. v. Excel Communications, Inc.*, 172 F.3d 1352, 1358. (Fed. Cir. 1999) (Emphasis added); *See State Street Bank & Trust Co. v. Signature Fin. Group, Inc.*, 149 F.3d 1368, 1373, 47 USPQ2d 1596, 1601 (Fed.Cir.1998). The inquiry into patentability requires an examination of the contested claims to see if the claimed subject matter, as a whole, is a disembodied mathematical concept representing nothing more than a "law of nature" or an "abstract idea," or if the mathematical concept has been ***reduced to some practical application rendering it "useful."*** *AT&T* at 1357 citing *In re Alappat*, 33 F.3d 1526, 31 1544, 31 U.S.P.Q.2D (BNA) 1545, 1557 (Fed. Cir. 1994) (emphasis added).

Independent claim 1 (and similarly independent claims 19, 30, 42, 44, 53, 54, 62-64) recites a computer implemented system that facilitates building a statistical model for a computer readable data set, comprising *a first training algorithm that efficiently builds a rough statistical model from a subset of the computer readable data set capable of statistical characterization, an evaluation component that evaluates the rough statistical model to determine whether the subset of the computer readable data set is an appropriate subset to build a statistical model for the computer readable data set, a*

*second training algorithm that builds a refined statistical model for the computer readable data set from the subset if deemed appropriate by the evaluation component, the refined statistical model discovers good clustering of data for a fixed number of clusters and a data scheduler that, based on a data policy, adaptively controls the size of subsets for which the first training algorithm is applied to facilitate building a more accurate statistical model.* In the subject Office Action, it is contended that a useful, concrete and tangible result is not found because the claimed invention lacks a practical application. Applicants' representative respectfully disagrees. Claim 1 (and similarly independent claims 19, 30, 42, 44, 53, 54, 62-64) recites acts that construct a refined statistical model from statistically characterizable data in a computationally economic way. The refined model enables further statistical analysis such as, *inter alia*, effective data clustering, statistical hypothesizing, and statistical prediction. Thus, such acts produce a useful, concrete and tangible result, namely the computationally efficient construction, from a data set, of a statistical model employable to statistical analysis of the characterized data set features.

Furthermore, in the subject Office Action, it is contended that the claimed invention merely manipulates abstract ideas. Applicants' representative avers to the contrary. The claimed subject does not manipulate abstract data but rather builds a useful statistical model from statistically characterizable data. Further, the constructed statistical model is employable to further analysis (*e.g.*, data clustering) of the underlying population or phenomena, which the computer readable data set represents and the model characterizes.

Applicants' claimed invention produces a useful, concrete and tangible result. Accordingly, withdrawal of this rejection is requested.

## **II. Rejection of Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60 and 62-64 Under 35 U.S.C. §112**

Claims 1, 3-25, 27-30, 32-42, 44-48, 50-60 and 62-64 stand rejected under 35 U.S.C. §112, first paragraph because current case law (and accordingly, the MPEP) require such rejection for claims that stand rejected under 35 U.S.C. §101. This rejection should be withdrawn for at least the following reasons. As stated above, independent

claims 1, 19, 30, 42, 44, 53, 54, and 62-64 have been amended to further emphasize aspects of the claimed subject matter. Accordingly, this rejection should be withdrawn.

### III. **Rejection of Claims 1, 19, 30, 42, and 64 Under 35 U.S.C. §102(b)**

Claims 1, 19, 30, 42, and 64 stand rejected under 35 U.S.C. §102(b) as being anticipated by Guha, *et al.* (US 5,140,530). Withdrawal of this rejection is requested for at least the following reasons. Guha *et al.* does not disclose or suggest all limitations set forth in the subject claims.

For a prior art reference to anticipate, 35 U.S.C. §102 requires that “*each and every element* as set forth in the claim is found, either expressly or inherently described, in a single prior art reference.” *In re Robertson*, 169 F.3d 743, 745, 49 USPQ2d 1949, 1950 (Fed. Cir. 1999) (*quoting Verdegaal Bros., Inc. v. Union Oil Co.*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987)) (emphasis added).

Applicants’ claimed subject matter relates to a system and method to facilitate building a model to characterize data based on a subset of the data having an appropriate size. The claimed subject matter constructs a crude model for an initial subset of data using a first parameter estimation algorithm. The model may be evaluated, for example, by applying the model relative to a holdout data set of the data. Where the model is unacceptable, additional data can be added to the data subset and the first parameter estimation algorithm is repeated for the aggregate data subset. An appropriate subset of the data exists when the first parameter estimation algorithm produces an acceptable model. The appropriate subset of the data subsequently can be employed by a different parameter estimation algorithm to build a statistical model that more accurately characterizes the data in its entirety. The subject matter as claimed in one aspect provides a relatively fast determination of an adequate size for the training data in situations where parameters will be estimated by employing a known parameter estimation technique (*e.g.*, an Expectation Maximization (EM) algorithm, and the like). To this end, independent claim 1 recites: ***a data scheduler that, based on a data policy, adaptively controls the size of subsets for which the first algorithm is applied***, independent claims 19, 30, 42,

and 64 recite: *utilizes a stopping criterion that is functionally related to an expected incremental benefit and an expected incremental cost associated with increasing the size of the subset*. Guha *et al.* does not disclose or suggest these aspects of the claimed subject matter.

Rather, Guha *et al.* relates to genetic learning techniques to evolve neural network architectures for applications where a general representation of neural network architecture is linked with a genetic learning strategy creating an environment for the construction of custom neural networks. In particular, the cited document involves the use of genetic algorithm methods to design new neural networks. The genetic algorithm (GA) is a robust function optimization method used in preference to gradient descent techniques for problems fraught with local minima, discontinuity, noise, or large numbers of dimensions. However, Guha *et al.* does not disclose adaptively controlling the size of the subset for which the first algorithm is applied. While Guha *et al.* discloses utilizing a set of input-output examples referred to as the training set for supervised learning (See col. 3, ll. 10-45), Guha *et al.* is silent regarding adaptively controlling the size of the training set. The training set is utilized to evaluate a constructed neural network and the cited reference nowhere discloses modifying or adapting the size of the training set.

Moreover, Guha *et al.* does not disclose, teach or suggest a stopping criterion that is functionally related to an expected incremental benefit and an expected incremental cost associated with increasing the size of the subset. Guha *et al.* relates to employing a genetic algorithm to create neural networks. The genetic algorithm are updated based on a fitness of a neural network. The fitness of a network is a measure of its worth on a problem, taking into account learning speed, accuracy and cost factors such as the size and complexity of the network. (See col. 2, ln. 63 to col. 3, ln. 2). While Guha *et al.* discloses the evaluation of neural network that weighs benefits and costs, this evaluation is employed as a comparison to other neural networks and not as a stopping criterion utilized to terminate incremental increases in the size of the subset of data. For example, the size and complexity of a neural network in Guha *et al.* is considered to determine a fitness of a neural network but is not altered by the fitness evaluation. A neural network that solves a problem just as well as a larger and/or more complex neural network may be deemed more fit. Still, Guha *et al.* is silent regarding altering the size of a neural

network or making a decision to cease further size increases based upon a fitness evaluation. Thus, Guha *et al.* fails to disclose, teach or suggest this aspect of the claimed subject matter.

In view of at least the foregoing, it is respectfully submitted that Guha *et al.* does not disclose, teach or suggest applicants' claimed invention as recited in independent claims 1, 19, 30, 42 and 64. Accordingly, this rejection should be withdrawn and the claims allowed.

**CONCLUSION**

The present application is believed to be in condition for allowance in view of the above comments and amendments. A prompt action to such end is earnestly solicited.

In the event any fees are due in connection with this document, the Commissioner is authorized to charge those fees to Deposit Account No. 50-1063[MSFTP184US].

Should the Examiner believe a telephone interview would be helpful to expedite favorable prosecution, the Examiner is invited to contact applicants' undersigned representative at the telephone number below.

Respectfully submitted,  
Amin, Turocy & Calvin, LLP

/Himanshu S. Amin/  
Himanshu S. Amin  
Reg. No. 40,894

Amin, Turocy & Calvin, LLP  
24<sup>TH</sup> Floor, National City Center  
1900 E. 9<sup>TH</sup> Street  
Cleveland, Ohio 44114  
Telephone (216) 696-8730  
Facsimile (216) 696-8731